Matrix tra scienza e filosofia[*]

Marco Salucci

Società Filosofica Italiana e Liceo Scientifico Gramsci, Firenze

Andrea Sani

Società Filosofica Italiana e Liceo Classico Galileo, Firenze

1. Il ciclo di *Matrix*

Nel 1999 esce il film di fantascienza *Matrix* (1999) di Andy e Larry Wachowski, cui sono seguiti, nel 2003, *Matrix Reloaded* e *Matrix Revolutions*, due ulteriori capitoli della saga molto inferiori all'originale.

La trama dell'intero ciclo è notissima. Neo (Keanu Reeves), un abile pirata informatico, viene contattato da un gruppo di *hackers* accusati di terrorismo. La misteriosa Trinity (Carrie Ann Moss) lo conduce da Morpheus (Laurence Fishburne), capo del gruppo di hackers ribelli. Neo apprende così che la sua vita, come quella di tutta l'umanità, è un'illusione prodotta dalle macchine, che utilizzano gli esseri umani per ricavarne energia. Nel 2099 le macchine hanno vinto la guerra contro gli uomini e li hanno imprigionati in capsule che contengono una sorta di liquido amniotico. I prigionieri dormono un sonno artificiale, mentre i loro recettori sensoriali sono collegati a un gigantesco computer che, grazie a un programma di 'neurosimulazione interattiva', la Matrice, fa vivere gli esseri umani in una specie di sogno continuo, per cui tutte le sensazioni che provano (vista, udito, gusto, olfatto e tatto) risultano simulate. La Matrice proietta nel cervello dei prigionieri l'illusione di vivere nell'anno 1999 in un mondo reale, mentre essi sono corpi dormienti nelle vasche. Scoperta la verità, sarà proprio Neo, l'*eletto* destinato a liberare l'umanità, a dare battaglia alle macchine.

Matrix Reloaded e *Matrix Revolutions* propongono nuove rivelazioni, che però complicano in modo artificioso il messaggio espresso dal capitolo iniziale della saga. In *Matrix Reloaded*, le macchine attaccano la città sotterranea di Zion, l'ultimo rifugio degli esseri umani sulla Terra dopo il cataclisma nucleare. Nel corso di questo secondo capitolo della saga, apprendiamo che anche la ribellione di Neo è prevista e controllata da Matrix. La rivelazione è comunicata a Neo dall'Architetto (Helmut Bakaitis), il programma (con sembianze umane) che ha creato e che governa Matrix. L'Architetto svela che il 99% degli umani non si rende conto di vivere in un universo artificiale.

* Lezione tenuta a Firenze il 28 ottobre 2009, nell'Altana di Palazzo Strozzi, nell'ambito dell'edizione 2009 di *Pianeta Galileo*.

Esiste, però, un 1% che si 'risveglia' e che rifiuta la dimensione illusoria imposta dalle macchine. Per risolvere il problema le macchine hanno ideato la figura dell'eletto e la città di Zion (un nome che i fratelli Wachowski hanno ripreso dalla Bibbia, dove Zion è l'appellativo di Gerusalemme). Neo dice all'Architetto:

> Tu sei il risultato finale di una anomalia che nonostante i miei sforzi sono stato incapace di eliminare da quella che altrimenti è un'armonia di precisione matematica. Sebbene resti un problema costantemente arginato essa non è imprevedibile e pertanto non sfugge a quelle misure di controllo che ti hanno condotto inesorabilmente qui.

In *Matrix Revolutions* si svolge lo scontro finale tra Neo e l'Agente Smith (Hugo Weaving), il programma di difesa appositamente creato dalle macchine per eliminare qualsiasi minaccia al sistema grazie ai suoi incredibili superpoteri. Alla fine del primo film della trilogia, Neo crede di aver ucciso definitivamente Smith, disintegrandolo. In realtà, Smith stesso si è ribellato al controllo delle macchine, che vorrebbero 'cancellarlo' poiché ha fallito la sua missione. Smith si moltiplica a dismisura come virus informatico e la sua presenza costituisce ormai un problema per le stesse macchine. Al termine del terzo film, Neo, che assume sempre di più la funzione del Salvatore dell'umanità, lo sconfigge definitivamente ma muore nello scontro, sacrificandosi.

La scomparsa di Neo e di Smith porta alla pace fra gli umani e le macchine, che interrompono il devastante attacco che hanno mosso alla città di Zion.

2. Il mito della caverna

Il ciclo di *Matrix* si segnala per prolungate scene da *action movie* destinate a emozionare le masse dei giovani spettatori. Tuttavia, soprattutto il primo film della serie è anche ricchissimo di riferimenti filosofici, ampiamente sviscerati nel volume a cura di William Irwin, *Pillole rosse. Matrix e la filosofia* [7, vedi anche 4, 15]. Gli argomenti filosofici più significativi del film riguardano la distinzione fra apparenza e realtà e il problema del rapporto mente-corpo.

La problematica fondamentale di *Matrix* è l'opposizione fra mondo vero e mondo fittizio. Nel film si sostiene che la realtà è diversa da come ci appare, e si fa l'ipotesi che ci sia qualcuno che appositamente crea un mondo illusorio allo scopo di ingannarci e di tenerci sottomessi. Per questi temi, secondo molti commentatori[1] il film dei fratelli Wachowski può essere letto come un rifacimento contemporaneo del mito della caverna, contenuto nel VII libro della *Repubblica* di Platone e riguardante, appunto, il rapporto realtà-finzione.

Nel racconto mitico, Platone immagina che alcuni schiavi siano incatenati, sino dall'infanzia, in una caverna sotterranea e che siano costretti a vedere solo il fondo della grotta. Su questo fondo vengono proiettate le ombre di piccole statue che sporgono al di sopra di un muricciolo situato alle spalle dei prigionieri. I prigionieri, che non hanno mai visto altro, guardano le ombre e pensano che queste siano la realtà. In un secondo momento, uno dei prigionieri riesce a liberarsi, si volta e vede le statue che proiettano

le ombre, scoprendo che esse sono più reali di queste. Platone immagina, poi, che il prigioniero possa salire su per la via che porta all'apertura della caverna ed esca fuori conoscendo, finalmente, il mondo veramente reale. Lo schiavo vorrebbe rimanere per sempre alla superficie, senza ritornare nella caverna. Infatti, se egli tornasse sotto terra per informare i suoi compagni ancora schiavi di ciò che ha visto, andrebbe incontro a gravi pericoli: i suoi occhi, non più abituati all'oscurità, non saprebbero più vedere le ombre; perciò sarebbe deriso dai suoi compagni. Costoro, infine, infastiditi dal suo tentativo di scioglierli e portarli fuori dalla caverna, lo ucciderebbero [16, VII, 514 a-517 a].

In quest'ultima parte del mito platonico è simboleggiata la sorte del sapiente, che ha conosciuto la verità (il Mondo delle Idee, o Iperuranio) e ha il dovere di comunicarla agli altri uomini. Costui, però, corre il rischio di essere scambiato per pazzo da chi è condizionato dai pregiudizi. L'uccisione dello schiavo che si è liberato adombra dunque la sorte toccata a Socrate, il filosofo che fu condannato a morte dagli stessi cittadini ateniesi che egli intendeva liberare dall'ignoranza.

Nel film *Matrix*, Neo è un novello Socrate: ha scoperto la verità sepolta nella Matrice e vuole svelare agli altri uomini l'inganno di un'esistenza illusoria, fatta solo di 'ombre', cioè di immagini virtuali. Ma anche lui incontrerà degli uomini che non vogliono essere liberati e che cercheranno di ucciderlo, pur di non essere 'svegliati' dal sogno apparentemente confortevole imposto dalle macchine. Cypher (Joe Pantoliano), un compagno di Morpheus, rappresenta questa umanità che preferisce rimanere ignorante e schiava delle apparenze, piuttosto che riacquistare la propria libertà. Cypher tradisce Morpheus e passa dalla parte dell'Agente Smith, al quale rivela i piani dei suoi compagni. In cambio ottiene da Smith la possibilità di diventare un attore famoso nella dimensione illusoria di Matrix, dimenticando completamente le sue esperienze precedenti e l'esistenza del mondo reale.

3. Cervelli in una vasca

In *Matrix* gli esseri umani vivono come corpi incapsulati in vasche. Però, la storia raccontata dal film funzionerebbe anche se invece di corpi si trattasse solo di cervelli: le stimolazioni sensoriali indotte da Matrix hanno di fatto bisogno solo del sistema nervoso. Potremmo dunque immaginare che nelle vasche di Matrix non si trovino corpi ma cervelli. Non si tratterebbe in verità di una fantasia originale, dal momento che nella letteratura e nella filmografia fantascientifica si trovano innumerevoli esempi di storie in cui compaiono cervelli separati dai corpi, nutriti in vasche, che vivono senza un corpo, resuscitati o trapiantati in corpi diversi da quello originario. Citando un po' a caso – e senza la pretesa di scegliere le opere esteticamente più valide – ricordiamo il film *Il cervello che non voleva morire* (*The Brain that wouldn't die*, 1962) di Joseph Green, gli innumerevoli film ispirati a *Frankenstein*, lo sceneggiato Rai *Gamma* di Salvatore Nocita, andato in onda nel 1975, e *The atomic Brain* (USA, 1964) di Joseph Mascelli. Anche nei fumetti si trovano molte storie ispirate al tema in questione: in *Superman* di

Jerry Siegel e Joe Shuster (più precisamente con la comparsa del supercriminale *Ultra-Humanite*, introdotto nella serie nel 1939); nel manga giapponese *Black Jack* (1974) di Osamu Tezuka, o nella saga dell'*Incal* (1981-88) scritta da Alejandro Jodorowsky e disegnata da Moebius (alias Jean Giraud).

Il tema, indipendentemente dalla forma espressiva con cui viene sviluppato, è senza dubbio affascinante e filosoficamente molto stimolante. Una delle componenti di tale fascino consiste nel singolare ribaltamento della logica dei trapianti ordinari: mentre nei comuni trapianti d'organo, per esempio di cuore, è il donatore che muore, nei trapianti di cervello sarebbe invece il donatore a sopravvivere. Dunque alla seduzione della sopravvivenza dopo la morte si aggiunge quella di vivere una vita diversa, in un altro corpo, in un altra situazione, con altre *chances*. Magari chi è nato ciabattino può sperare di vivere da principe, come immagina nel XVII secolo il filosofo inglese John Locke cercando di districarsi nella difficile questione circa cosa sia a determinare l'identità personale [11, cap. XXVII, § 17][2].

Il problema del rapporto fra apparenza e realtà è già ampiamente discusso nella storia della filosofia sin dall'epoca di Parmenide. Tuttavia, l'immagine del cervello nella vasca ha origine nella letteratura fantascientifica. Forse si tratta anzi di uno di quei non infrequenti casi in cui la fantasia ha preceduto la filosofia, e lo conferma il fatto che proprio il filosofo che più di ogni altro ha influenzato la discussione contemporanea al riguardo sia, per sua esplicita dichiarazione, appassionato lettore di fantascienza fin dalla giovinezza[3]. Alludiamo a Hilary Putnam che in un capitolo del suo libro *Ragione verità e storia* [18] sfida il lettore a dimostrare di non essere un cervello in una vasca che si illude di leggere un capitolo del libro di Putnam *Ragione verità e storia* opportunamente stimolato da un computer che gli fa credere di leggere un capitolo del libro di Putnam *Ragione verità e storia*[4].

Come si può immaginare il compito non è facile. Si tratta di una classica sfida scettica, la stessa sfida che il filosofo Cartesio, nel XVII secolo, aveva affrontava quando scriveva:

> Che vedo io da questa finestra, se non dei cappelli e dei mantelli, che potrebbero coprire degli spettri o degli uomini finti, mossi solo per mezzo di molle?[5]

Cartesio ipotizza che un genio maligno lo stia ingannando sistematicamente e che gli faccia credere erroneamente di avere un corpo, di vivere in un mondo reale e persino, erroneamente, che 2+2=4. Non potrebbe questo genio maligno essere l'Architetto di *Matrix*?

La plausibilità di immaginazioni come quelle proposte da Cartesio o da Putnam è confermata dalla scienza stessa. I primi esperimenti in tal senso risalgono alla metà del XX secolo, quando il neurochirurgo canadese Wilder Penfield cominciò a sperimentare un trattamento chirurgico di casi gravi di epilessia. Durante l'intervento, Penfield stimolava con microelettrodi il cervello dei pazienti inducendo in essi l'illusione di vivere realmente un certo un evento quotidiano. Mentre stimolava varie aree della corteccia

cerebrale, Penfield chiedeva ai pazienti – in anestesia locale e coscienti – che cosa provassero ottenendone risposte come: «Ho udito qualcosa, una madre che chiamava il suo bambino. È qualcuno del vicinato, dove abito io», oppure: «Sento delle voci. Laggiù, lungo il fiume. Sono un uomo e una donna. Vedo il fiume», e ancora: «Sono in ufficio, un collega appoggiato a una scrivania mi chiama, ha in mano una matita» [12, p. 47].

4. Il problema del rapporto mente-corpo

Si noti che quando Cartesio immagina di essere ingannato da un genio maligno – la stessa situazione rappresentata in *Matrix* dagli uomini che pur vivendo in vasche credono di vivere una vita reale – e quando immagina che i suoi simili siano automi – una situazione che invece non è presente in *Matrix* ma che lo è in moltissimi altri film o racconti di fantascienza: uno per tutti *A.I. Intelligenza artificiale* (*A.I. Artificial Intelligence*, 2001) di Steven Spielberg – sta in realtà facendo un'unica ipotesi: quella della separabilità del corpo dalla mente. La mente non è il cervello: è questa la tesi fondamentale sostenuta da Cartesio e nota come *dualismo* proprio poiché ritiene che mente e corpo siano *due* sostanze diverse capaci di esistenza autonoma e di interazione[6].

Sia chiaro: nessuno può sostenere ragionevolmente che la mente sia identica al corpo; sarebbe come sostenere che il movimento è identico ad avere un paio di gambe, con l'indesiderabile conseguenza che senza un paio di gambe non ci sarebbe movimento: quadrupedi, automobili, aerei e navi stanno a dimostrare il contrario. Il punto è un altro: può esistere – per continuare con la metafora – il movimento senza un qualsivoglia apparato fisico (gambe, ruote, eliche)? Il dualismo sostiene che la mente è appunto in grado di esistere anche senza un apparato fisico. Il fatto che la mente sia incarnata in un corpo è accidentale per il dualismo: mente e corpo sono ontologicamente distinti.

I dualisti, e Cartesio fra questi, hanno prodotto molti argomenti a sostegno della loro tesi; per rimanere fedeli alle consegne di questo intervento ne ricorderemo solo uno reperibile sia nella letteratura filosofica che nella filmografia: il film *Viaggio allucinante* (*Fantastic Voyage*, 1966), di Richard Fleischer, tratto da una storia di Jay Lewis Bixby e Otto Klement (e da cui Isaac Asimov ha poi ricavato un romanzo). Nella pellicola si immagina che un sottomarino miniaturizzato viaggi dentro un corpo umano. Giunto nel cervello del malcapitato, un membro dell'equipaggio del sottomarino nota che si vedono solo scariche elettriche percorrere le connessioni fra i neuroni ma non si vedono pensieri o ricordi.

Anche se quasi sicuramente non risulta tra le fonti del film di Fleischer, la *Monadologia* (1714) del filosofo tedesco Gottfried Wilhelm Leibniz descrive una situazione molto simile a quella di *Viaggio allucinante*; solo che per Leibniz, all'inizio del Settecento, i fenomeni che avvenivano nel corpo e nel cervello erano di natura meccanica, non elettrica, come quelli di un mulino[7]. La sostanza dell'argomento di Leibniz è questa: se i fenomeni mentali fossero fisici allora potremmo osservarli; ma non li osserviamo, dunque non sono fisici [10, § 14 della *Monadologia*].

Molti argomenti utilizzati a favore del dualismo sono riducibili a quest'argomen-

to di Leibniz, ed anche se si dimentica di citarne la vera paternità e la si attribuisce a un paio di filosofi nostri contemporanei[8], resta alla base di quello che oggi si chiama *Argomento della conoscenza*. Anzi, è ancora Leibniz che ha fornito gli strumenti per orientarsi in modo sufficientemente chiaro fra i numerosi argomenti che sono stati forniti dai dualisti a sostegno della loro tesi. La cosiddetta *Legge di Leibniz* esige che se due entità sono uguali allora devono avere le stesse proprietà, se hanno proprietà diverse non sono uguali. Così scopriamo che l'argomento della conoscenza e quello del mulino poggiano sul presupposto che se il sistema nervoso ha la proprietà di essere osservabile e la mente no, allora non possono essere la stessa cosa. Ma ovviamente potremmo trovare anche altre proprietà possedute da uno solo dei due elementi della coppia mente-corpo; per esempio – è ancora Cartesio a sostenerlo – l'estensione: il mio cervello occupa uno spazio, il mio pensiero no (abbiamo detto «occupa uno spazio», non «si trova in uno spazio»: ovviamente il mio pensiero si trova nella mia testa). Tutto bene allora per il dualismo? Beh, non proprio.

Il modello cartesiano dell'uomo è stata definito da un filosofo americano, Gilbert Ryle, il *modello del fantasma nella macchina*: l'anima abita nel corpo come un fantasma in un robot e lo guida come i soldati dell'impero guidano i bipedi *land walker robot* del ciclo cinematografico di *Star Wars* (1999-2005) di George Lucas. Le interazioni che avvengono fra l'anima e il corpo vanno, per così dire, in due direzioni: o dal corpo verso l'anima, come nelle sensazioni, o dall'anima verso il corpo, come nei movimenti. Ma come possono interagire anima e corpo se l'anima è immateriale? Come può avvenire il contatto fra le due? Non accadrebbe ciò che accade a Ulisse nel libro XI dell'*Odissea* quando nell'Ade incontra l'anima della madre e cerca di abbracciarla? «Tre volte dalle mie mani, all'ombra simile o al sogno, volò via». Elisabetta Stuart, ex regina di Boemia ed esule nei Paesi Bassi durante la Guerra dei Trent'anni, pose il problema a Cartesio in un paio di lettere che costituiscono, insieme alle risposte di Cartesio, la prima chiara formulazione del problema del dualismo, problema noto come *Problema dell'interazione.*

Cartesio ed Elisabetta pensavano che le interazioni potessero avvenire solo per contatto e questo rendeva particolarmente acuto il problema dell'interazione mente-corpo. E tuttavia, secondo i critici del dualismo, anche se si ipotizzasse un'interazione mente-corpo senza contatto, una sorta di induzione invece che elettromagnetica psicofisica, le cose non andrebbero meglio per i cartesiani. L'interazione con una sostanza spirituale – dove per *spirituale* intendiamo non materiale – viola un principio utilizzato dalle scienze naturali: quello della chiusura causale del mondo naturale. Secondo tale principio, gli eventi fisici possono avere solo cause fisiche; pertanto se un movimento del corpo (come il movimento di una mano) ha per causa uno stato mentale (la volontà di muovere la mano), lo stato mentale dovrebbe essere fisico.

I sostenitori del dualismo interazionista potrebbero ribattere che il principio della chiusura causale è un'assunzione metodologica valida per i fisici, ma che non può trasformarsi in un principio metafisico, valido in assoluto circa il mondo in tutti i suoi

aspetti. Anche se in molti casi abbiamo trovato finora una causa fisica degli eventi, non possiamo logicamente concludere che tutti i fenomeni abbiano una causa fisica sufficiente. L'induzione non può giustificare il passaggio da "molti" a "tutti". Ne deriva, secondo questa impostazione, che non si può escludere a priori che un fatto si sottragga a tale chiusura causale e che quindi abbia un causa mentale. Va comunque precisato che la negazione del principio per cui eventi fisici possono essere causati solo da eventi fisici, comporta anche la limitazione della validità di un altro principio della fisica, e cioè del principio di conservazione dell'energia (primo principio della termodinamica), di cui la chiusura del mondo fisico è una diretta conseguenza.

D'altra parte, se l'ipotesi cartesiana non è esente da critiche, anche contro la teoria opposta, e cioè contro la concezione materialista che identifica gli stati mentali con gli stati fisici, non mancano gli argomenti (quelli che abbiamo già citato e che sono utilizzati dai dualisti).

5. Il funzionalismo

Fin qui, dunque, la discussione somiglia a una partita di ping-pong: le obiezioni e le risposte rimbalzano da una parte all'altra senza che nessun giocatore possa vantare il possesso definitivo della pallina. In realtà, a chi conosce la letteratura sull'argomento la metafora del ping-pong apparirà risibile: la partita che si gioca sul campo del problema mente corpo più che una partita a due sembra un'olimpiade: centinaia di atleti e decine di specialità [20, 21].

Per la verità, la prospettiva che riduce gli stati mentali a stati cerebrali non è neppure esplicitamente presente in *Matrix*. Nel film dei fratelli Wachowski non c'è nessun Dr. Manhattan della miniserie a fumetti *Watchmen* (di Alan Moore e Dave Gibson, 1986-87), che affermi un riduzionismo radicale come quando dichiara di non essere turbato dalla morte del suo ex collega *Il Comico*, perché «un corpo vivo e un corpo morto contengono lo stesso numero di particelle, strutturalmente non c'è nessuna differenza». Piuttosto, accanto a menti che vivono separate dal corpo in una realtà illusoria, il mondo di *Matrix* è popolato da programmi. Le menti sono programmi: lo sono l'Architetto, l'Oracolo, l'Agente Smith. *Matrix* è una saga funzionalista, dove per *funzionalismo* si intende la teoria secondo la quale le menti, anche quelle umane, sono programmi. È quanto accade anche nel film *Il tredicesimo piano* (*The Thirteenth Floor*, 1999) di Josef Rusnak, nel quale il protagonista scopre di essere un programma e di vivere in un mondo virtuale, come un personaggio in un videogioco. Con ciò il catalogo delle possibilità è completato: in *Matrix* ci sono ancora esseri umani, benché illusi da programmi di vivere un vita reale, in *Tredicesimo piano,* in *Matrix Reloaded* e in *Matrix Revolutions* l'universo sembra popolato solo da programmi: dagli esseri umani che sognano una vita reale siamo passati ai programmi che sognano di essere esseri umani.

È il solito Putnam ad aver proposto, negli anni Sessanta del secolo scorso, l'idea secondo la quale la mente è un *software*; all'epoca, prima che egli mutasse posizione, il suo motto era: «la mente sta al *software* come il cervello all'*hardware*». La prospettiva

funzionalista propugnata da Putnam fornisce non solo una risposta alternativa a quelle dualista e materialista, ma pone proprio una domanda diversa: non *cos'è* la mente – spirito o materia – ma *cosa fa*. In molte occasioni questa è la sola prospettiva giusta. Se, per esempio, volete definire una pompa non potete dire che una pompa è un meccanismo fatto di un cilindro e di un pistone, facendo così riferimento ai componenti fisici, ma dovete dire che una pompa è un meccanismo qualsiasi atto a spostare liquidi. Tale meccanismo potrebbe essere fatto allora non solo con pistoni e cilindri, ma anche con ingranaggi (le pompe della benzina delle auto), con eccentrici (le pompe per la circolazione sanguigna extracorporea). Sciolta la metafora, non è il cervello che conta ma il *software* che ci 'gira' dentro. Lo sviluppo dell'Intelligenza Artificiale e l'affermazione del funzionalismo si sono incontrati e hanno dato vita a una scuola psicologica che è oggi dominante: il cognitivismo. Secondo il cognitivismo l'uomo è, appunto, un elaboratore d'informazioni.

> La nostra sostanza, ciò di cui siamo fatti, non stabilisce pressoché nessuna restrizione [...] alla nostra forma, e ciò che veramente ci interessa è la forma, non la materia. La domanda da porsi non è: che cos'è la materia? bensì *che cos'è la nostra forma intellettuale?* E qualunque cosa sia la nostra sostanza, spirito, materia, o emmenthal svizzero, essa non porrà alcuna interessante restrizione [...] alla risposta a tale domanda [17, p. 330].

Il principale vantaggio di tale prospettiva sta, secondo i funzionalisti, nell'evitare l'antropocentrismo consistente nell'attribuire stati mentali solo a esseri dotati di sistema nervoso: perché non potrebbero avere stati mentali i marziani, i computer o i robot del futuro? Insomma tutti quegli esseri che popolano il variegato mondo della letteratura e della cinematografia di fantascienza? È un dato di fatto che gli animali, uomo compreso, abbiano un cervello, ma gli stati mentali che essi hanno potrebbero funzionare altrettanto bene su 'macchine' fisicamente diverse: computer, robot, alieni.

Ora, non vorremmo negare che i marziani e i robot possano avere stati mentali; quello che non sembra plausibile è che marziani e robot abbiano stati mentali come i nostri. Funzionalismo e cognitivismo non tengono in nessun conto gli aspetti fisici dell'entità in cui il programma è realizzato, quindi non tengono in nessun conto gli aspetti biologici degli unici esseri che, almeno attualmente, sappiamo possedere stati mentali. Il funzionalismo è dunque il tentativo di elaborare una psicologia universale valida per ogni essere reale o possibile indipendentemente da come è fisicamente costituito l'organismo e l'ambiente in cui vive. Molti dati dell'anatomia e della fisiologia comparate rendono tuttavia tale pretesa assai problematica. Per limitarci al caso della percezione, le modalità di percezione delle specie animali sono così varie che rendono difficile parlare di stati mentali transpecifici. Come è possibile sostenere che i pesci abbiano, per esempio circa l'acqua, gli stessi stati mentali degli uccelli? O che i nostri gusti e quelli dei pipistrelli sugli insetti siano gli stessi? Gli stati mentali, lungi dall'essere indipendenti dal materiale in cui 'girano' sembrano essere dipendenti anche da molti fattori biologici, fisiologici, percettivi, motori, ecologici ecc[9].

Le caratteristiche ambientali alle quali i vari organismi devono mostrarsi sensibili variano a seconda delle loro necessità biologiche. Le capacità percettive dunque si sono evolute in modo da selezionare le caratteristiche salienti. A questo scopo gli organi di senso presentano rivelatori di caratteristiche diversi per ogni specie. L'esistenza di rivelatori di caratteristiche specie-specifici è un forte argomento a sostegno dell'esistenza di stati mentali specie-specifici.

Al di là delle differenze siamo ora in grado di dire che c'è un'affinità fra dualismo e funzionalismo: entrambi accentuano la separatezza fra mente e corpo, anche se tale separatezza assume nella prima teoria la forma dell'indipendenza e nella seconda quella dell'indifferenza.

Recentemente è emersa – o meglio, è riemersa – una nuova prospettiva circa il rapporto mente-corpo: quella cosiddetta della *mente incorporata*. I lavori del neuroscienziato Antonio Damasio sono stati fra i primi a segnare l'inizio della nuova tendenza; egli denuncia 'l'errore cartesiano' di separare la mente dal corpo e utilizza i risultati delle ricerche relative al sistema nervoso centrale per argomentare a favore della tesi per la quale la mente emergerebbe del corpo. Secondo Damasio, i processi di pensiero emergerebbero da una serie di mappe del corpo costruite dal cervello.

Senza risalire troppo indietro nel tempo[10], l'idea che gli stati mentali dipendano dal modo in cui la mente è *embodied* e *entangled* in e con il corpo è stata avanzata anche dal matematico Henri Poincaré. Egli lavorava a un problema: com'è possibile che percepiamo uno spazio tridimensionale se tutte le immagini sulla nostra retina sono soltanto bidimensionali?

> Ebbe l'idea di introdurre nella percezione il movimento: quando guardiamo non siamo oggetti stazionari, ma ci muoviamo. E così osservò che la complessità sufficiente per computare le nozioni di spazio, di oggetto, ecc. può prodursi soltanto quando si possono correlare i mutamenti percettivi associati con le azioni motorie. Abbiamo così bisogno di un accoppiamento di attività sensoriali e motorie, e ciò produce quello che vorrei chiamare l'anello sensomotorio: le attività sensoriali informano le attività motorie che a loro volta informano le attività sensoriali. Cinquant'anni dopo, le stesse cose furono più o meno il risultato delle ricerche di Jean Piaget. Egli fece notare come al bambino sia indispensabile l'attività senso-motoria per costruire qualunque idea di realtà. I bambini devono afferrare, allungare le mani, mettere in bocca, sentire, toccare, agitarsi. E solo allora possono vedere. Si vede con le mani o – come dice Maturana – si vede con i piedi, perché soltanto quando si cammina le cose cambiano, e si può così produrre l'anello senso-motorio[11].

I risultati delle neuroscienze sono oggi ritenuti quasi unanimemente indispensabili per capire aspetti essenziali e rilevanti della nostra realtà mentale. Tuttavia, sebbene le neuroscienze avanzino la pretesa di dire l'ultima parola circa la persona umana, non si può dimenticare che l'ipotesi di una completa riducibilità delle manifestazioni della coscienza al puro dato biologico è tutta da dimostrare. A una prospettiva legata esclusivamente ai processi neurobiologici sfugge, per esempio, l'aspetto interiore, privato e

qualitativo delle nostre esperienze. Alcuni filosofi della mente, i cosiddetti *misteriani* come Colin McGinn, ritengono che la scienza non possa rispondere a questioni che riguardano i *qualia*, la coscienza, il libero arbitrio, anche se i suoi progressi aiutano a definire e circoscrivere sempre meglio tali concetti e tali questioni[12].

6. L'armonia prestabilita

Circa il complesso problema del rapporto mente-corpo, *Matrix* sembra alludere a un'ulteriore, possibile, soluzione, proposta dal già citato filosofo tedesco Gottfried Wilhelm Leibniz, e che potremmo definire del *parallelismo*. Il parallelismo è un'ipotesi di stampo dualista, ma in contrapposizione all'interazionismo cartesiano.

Stranamente, Leibniz non è nemmeno citato nel libro su *Matrix* a cura di Irwin, *Pillole rosse*, anche se in un saggio di Slavoi Žižek contenuto nel volume si mette in relazione il film dei fratelli Wachowski con la filosofia occasionalista di Nicolas Malebranche[13], contemporaneo di Leibniz. Eppure sono assai numerose le affinità tra le ipotesi fantascientifiche di *Matrix* e le teorie leibniziane. Infatti, Leibniz è stato l'ideatore dell'aritmetica binaria, che consente di rappresentare tutti i numeri con due cifre soltanto, 0 e 1, e che gioca un ruolo importantissimo nel campo dell'informatica, dove qualsiasi messaggio è convertito in un alfabeto composto da due soli simboli – 0 e 1, appunto – detti *bit* (dall'inglese *BInary digiT*). Ricordiamo anche che Leibniz, per le sue ricerche di logica e per il suo tentativo di ridurre il ragionamento a un calcolo, è considerato il precursore dell'Intelligenza Artificiale, e cioè di quel campo di indagine che si pone l'obiettivo di riprodurre su una macchina alcune delle attività mentali dell'uomo. Inutile sottolineare che l'informatica e l'Intelligenza Artificiale stanno alla base della pellicola ideata dai fratelli Wachowski. Ma i collegamenti più significativi tra *Matrix* e il pensiero leibniziano riguardano più che altro il tema della cosiddetta «armonia prestabilita», una teoria dualista introdotta da Leibniz per spiegare in modo *non cartesiano* il rapporto fra la mente e il corpo. Secondo tale teoria, le due serie di eventi – quelli mentali e quelli fisici – corrono paralleli e sincronici, senza interagire. Ma vediamo la proposta di Leibniz un po' più nel dettaglio, collegandola alla sua visione del mondo espressa nella *Monadologia* del 1714 [10].

A giudizio di Leibniz, tutta la realtà è composta da 'monadi', ossia da centri di forza o atomi immateriali di pura energia. L'uomo, in particolare, è una sostanza formata da monadi inferiori, costituenti il corpo, e da una monade superiore, detta «monade dominante» o «spirito». La monade spirito, secondo Leibniz, non interagisce con le monadi del corpo. Tutte le percezioni della nostra anima riguardanti il mondo scaturiscono non dall'esterno, ma dal suo interno, spontaneamente, secondo le leggi della sua natura, dato che le monadi, essendo immateriali, sono senza «finestre», cioè senza aperture attraverso le quali qualcosa possa entrare o uscire [10, § 7]. Nella percezione, il nostro spirito non esce fuori di sé, ma si modifica in virtù di una forza rappresentativa interna a se stesso. Ne consegue che la vita dell'anima è come un sogno, anche se coerente e ben regolato.

Ma come si spiega, allora, l'apparente interazione causale che sembra sussistere fra la monade-spirito e le monadi del corpo negli atti volontari (quando l'anima sembra agire sul nostro organismo fisico), e nelle sensazioni (quando il corpo sembra agire sullo spirito)? E come si giustifica il rapporto che pare sussistere fra la monade-spirito di un individuo e quella di un altro individuo?

Per chiarire l'accordo fra le monadi, Leibniz ricorre, appunto, alla dottrina dell'armonia prestabilita. Secondo Leibniz, l'accordo fra le varie monadi è garantito da Dio grazie a un'armonia prestabilita sin dall'eternità tra le attività di tutte le sostanze, in virtù della quale, nell'istante preciso in cui una sostanza intende agire sull'altra, si verificano in quest'ultima le modificazioni corrispondenti. Secondo Leibniz, per esempio, quando la mia monade-spirito decide che il mio braccio si sollevi, il braccio si muove non a causa della mia volontà, ma perché le monadi del mio corpo sono sincronizzate da Dio a sollevare il braccio proprio quando la mia mente lo vuole. Il sollevarsi del braccio non è l'effetto della mia volontà di sollevarlo. Fra i due eventi non c'è un rapporto di causa ed effetto, ma solo un parallelismo frutto dell'armonia prestabilita.

Questo accordo è operato da Dio una volta per tutte, al momento della creazione. In base al «programma iniziale» stabilito da Dio, tutti gli stati successivi di ogni monade si susseguono l'uno all'altro sviluppandosi dalla natura stessa della sostanza individuale, senza che si verifichi alcun influsso esterno. Ma, contemporaneamente, tutti i suoi mutamenti corrispondono a quelli che avvengono in tutte le altre monadi che sussistono nel suo universo. La perfetta corrispondenza fra le monadi è paragonata da Leibniz alla sincronizzazione di due orologi che, pur essendo ciascuno indipendente dall'altro, segnano sempre la stessa ora. Afferma Leibniz:

> I corpi agiscono come se (per assurdo), le anime non ci fossero, e le anime agiscono come se non ci fossero i corpi, e tutti e due agiscono come se gli uni influissero sulle altre» [9, vol. VI, p. 595].

L'armonia prestabilita è però un'interpretazione molto strana del rapporto causale tra i vari esseri giacché, se tale ipotesi fosse vera, tutta la sapienza e tutta la potenza del Creatore sarebbero impiegate per darci l'illusione che le cose agiscano senza agire veramente! Questa concezione di Leibniz spinse il filosofo inglese Bertand Russell, nella *Prefazione* al suo saggio del 1900 *Esposizione critica della filosofia di Leibniz*, ad affermare che la *Monadologia*, a una prima lettura, gli era sembrata «una specie di racconto di fate»[14]. Se Russell avesse scritto il suo libro al giorno d'oggi, avrebbe detto probabilmente che il capolavoro leibniziano era paragonabile a 'una storia di fantascienza', tipo *Matrix*.

6. *Matrix* e la *Monadologia*

Singolari sono, infatti, le analogie tra *Matrix* e la metafisica leibniziana. Nel film dei fratelli Wachowski, i prigionieri delle macchine risultano, in un certo senso, delle monadi radicalmente isolate le une dalle altre. Il mondo virtuale in cui vivono i prigionieri, che è proiettato nelle loro menti da *Matrix*, è analogo alle percezioni delle sostanze

individuali di cui parla Leibniz, le quali non dipendono dall'azione della realtà esterna, ma scaturiscono direttamente dal loro interno, in virtù di un disegno divino.

Tuttavia, la Matrice non chiude ogni uomo in un suo universo artificiale individuale: infatti, coordina il programma di ciascuna persona con quello delle altre persone, in modo che tutta l'umanità viva il medesimo universo virtuale. Nel film dei fratelli Wachowski, grazie alla programmazione operata dal super-computer, ciò che accade a un individuo è attestato ed esperito da altri individui, pur non esistendo un rapporto diretto fra gli uomini. Analogamente, le varie monadi leibniziane vivono ognuna per proprio conto, ma sono sempre d'accordo tra loro, perché Dio ha coordinato le percezioni di tutte le sostanze individuali con l'armonia prestabilita. Cosicché, sia in *Matrix* sia nella *Monadologia*, il mondo è una sorta di sogno, che però risulta intersoggettivamente condiviso: anche i prigionieri di *Matrix* hanno sogni simili nello stesso momento.

Nell'universo del film, come in quello di Leibniz, la causalità diventa un'illusione radicale, per cui gli avvenimenti che si credono in rapporto di causa ed effetto non sono altro che avvenimenti in coincidenza per una preordinazione stabilita dall'alto. Per esempio, quando Neo pensa di compiere un movimento con il corpo, e il corpo – nella realtà virtuale – effettivamente lo esegue, il suo pensiero è 'causa' di tale moto non direttamente, ma solo in virtù di un programma di coordinamento stabilito da un super-computer.

Naturalmente, sussistono anche delle profonde differenze tra la situazione descritta da *Matrix* e quella ipotizzata dalla *Monadologia* leibniziana. Infatti, per Leibniz il Creatore dell'armonia prestabilita è un Dio perfettissimo e buono. Invece, nel pessimistico e cupo film dei fratelli Wachowski le macchine che hanno ideato la Matrice sono malvagie e sfruttano gli uomini, dai quali traggono l'energia necessaria per il loro funzionamento.

Ci sono altre differenze tra nell'universo di *Matrix* e la metafisica di Leibniz. Per esempio, in *Matrix* alcuni individui, fra i quali Neo, riescono a scollegarsi dalla Matrice (così, almeno, sembra nel primo film della trilogia); invece, le monadi, secondo Leibniz, non possono interrompere l'armonia prestabilita da Dio, e sono destinate, per tutta la loro esistenza, a rispettare la 'programmazione' fissata, una volta per tutte, all'inizio dei tempi.

I prigionieri della Matrice, poi, possono dirsi simili alle monadi solo perché sono chiusi nel loro mondo interiore e perché non interagiscono fra loro; tuttavia, non risultano certo, come le sostanze individuali leibniziane, dei puri spiriti privi di materia, dato che si tratta pur sempre di individui in carne e ossa.

Infine, tali prigionieri – anche se percepiscono uno spazio illusorio (il mondo del 1999) – risultano comunque collocati in uno spazio reale: infatti, nel 2099, sono disposti all'interno di vasche contigue che li collegano al super-computer. Al contrario, le monadi di Leibniz non esistono in alcuno spazio, in quanto, per il filosofo, spazio e tempo non sono reali ma ideali. Lo spazio – asserisce Leibniz – ben lungi dall'essere una

sostanza, non è neppure un ente. Esso è un ordine, come il tempo: lo spazio è l'ordine della coesistenza, mentre il tempo è l'ordine della successione tra gli eventi attestati dalle percezioni, e questi due tipi d'ordine esistono esclusivamente all'interno di ciascuna monade sotto forma di rappresentazione mentale. Proprio perché spazio e tempo sono relazioni, a giudizio di Leibniz non possono intercorrere tra una monade e l'altra, che non hanno alcun tipo di rapporto esterno fra loro.

A parte queste significative diversità, resta il fatto che *Matrix* può dirsi – utilizzando una terminologia proposta da Julio Cabrera nel suo libro *Da Aristotele a Spielberg. Capire la filosofia attraverso i film* (1999)[15]– un efficace 'concettimmagine' fantascientifico dell'armonia prestabilita leibniziana.

NOTE

[1] Vedi [23, pp. 144-147], Irwin, W., *Computer, caverne e oracoli: Neo e Socrate*, in [7, pp. 7-21], e Griswold Jr., C. L., *La felicità e la scelta di Cypher: l'ignoranza è un bene?,* in [7 pp. 153-167].

[2] Scrive Locke:

> Se l'anima di un principe portando con sé la consapevolezza della vita passata del principe, entrasse a informare di sé il corpo di un ciabattino subito dopo che questo fosse stato abbandonato della propria anima, ognuno vede che egli sarebbe la stessa persona che il principe, responsabile solo delle azioni del principe; ma chi direbbe che si tratta dello stesso uomo? [11, libro II, cap. XXVI, par. 17, , p. 400].

[3] A tal proposito, il confronto fra le date di pubblicazione dei romanzi e racconti di fantascienza e molti dei cosiddetti esperimenti mentali che si trovano nella filosofia anglosassone, crediamo, riserverebbe scoperte interessanti. Le pittoresche figure che popolano oggi la filosofia della mente – pipistrelli, neuro scienziati matti o vittime di esperimenti, zombi, cervelli nella vasca, alieni, robot, teletrasportati, superspartani ecc. – hanno abitato la letteratura fantastica prima di quella filosofica.

[4] Il tema del cervello nella vasca ha avuto, dopo il dibattito suscitato da Putnam, una risonanza tale che qualcuno si è messo a produrre T-shirt con su stampato «*I am a brain in a vat*», che possibile acquistare anche su internet.

[5] Descartes, R., *Meditazioni metafisiche, Prima meditazione* [1641], in [5, vol. I, pp. 199-204].

[6] Il dualismo è stata una delle teorie del rapporto mente-corpo fra le più criticate a causa del problema dell'interazione. Dualisti sono stati però, nel Novecento, il premio Nobel (1963) per la fisiologia e medicina John Carew Eccles e il filosofo Karl Popper, che hanno riproposto tale soluzione nel loro libro *L'Io e il Suo Cervello* del 1977 [6].

[7] Scrive Leibniz:

> Bisogna ammettere che la percezione e quello che ne dipende sono inesplicabili mediante ragioni meccaniche, cioè mediante figure e movimenti. Se immaginiamo una macchina costruita in modo che pensi, senta, percepisca, si potrà concepire che venga ingrandita conservando le medesime proporzioni, in modo che vi si possa entrare come in un mulino. Ciò fatto, nel visitarla internamente non si troverà altro che pezzi, i quali si spingono scambievolmente e mai alcuna cosa che possa spiegare una percezione. Cosicché questa bisogna cercarla nella sostanza semplice e non nel composto, o nella macchina [10, pp. 36-37].

[8] Thomas Nagel, Frank Cameron Jackson.

[9] Su questi temi vedi [24], [13], [14] e [8].

[10] Ma almeno per l'età moderna andrebbe citato Etienne Bonnot de Condillac.

[11] Foerster, H. V., *Cibernetica ed epistemologia: storia e prospettive*, in [1, p. 122].

[12] Per una discussione circa il problema del rapporto mente-corpo anche dal punto di vista della teologia, vedi [3]. Secondo O. Culmann le scritture cristiane insisterebbero più sulla resurrezione dei corpi che sull'immortalità dell'anima. Comunque Tommaso d'Aquino, il cui pensiero è alla base della concezione teologica cattolica, pur sostenendo che l'anima è aristotelicamente forma del corpo, e quindi intimante legata al corpo, argomenta nei suoi scritti che essa è una forma sussistente in grado di sopravvivere alla morte del corpo, e cioè che è immortale.

[13] Vedi Žižek, S., *Matrix, ovvero i due lati della perversione*, in [7, pp. 292-328].

[14] Confessa Russell: «Al pari di molti altri, sentivo che la *Monadologia* era una specie di fiaba fantastica [*a kind of fantastic fairy tale*], forse coerente ma del tutto arbitraria» [19, p. 31]. Russell ammette, però, di essersi ricreduto circa la plausibilità delle teorie leibniziane dopo la lettura del *Discorso di metafisica* del 1686 e della corrispondenza fra Leibniz e Antoine Arnauld: «Fu a questo punto che lessi il *Discours de métaphysique* e le lettere ad Arnauld. Un'ondata di luce si riversò improvvisamente in ogni più riposto recesso della costruzione filosofica di Leibniz».

[15] Vedi [2]. Sul rapporto tra il cinema e la filosofia vedi anche [23].

BIBLIOGRAFIA

[1] Bocchi, G., Ceruti, M. (a cura di), *La sfida della complessità*, Feltrinelli, Milano 1985.

[2] Cabrera, J., *Da Aristotele a Spielberg. Capire la filosofia attraverso i film* [1999], Bruno Mondadori, Milano 2000.

[3] Canobbio, G., *Il destino dell'anima. Elementi per una teologia*, Morcelliana, Brescia 2009.

[4] Cappuccio, M. (a cura di), *Dentro la Matrice. Filosofia, scienza e spiritualità in Matrix*, Edizioni Albo Versorio, Milano 2004.

[5] Descartes, R., *Opere*, 2 voll., Laterza, Roma-Bari 1967.

[6] Eccles, J. C., Popper, K. R., *L'Io e il Suo Cervello*, 3 voll., Armando, Roma 1981.

[7] Irwin, W., *Pillole rosse. Matrix e la filosofia*, Bompiani, Milano 2006.

[8] Lakoff, G., Nunez, R., *Da dove viene la matematica?*, Bollati Boringhieri, Torino 2005.

[9] Leibniz, G. W., *Die Philosophischen Schriften*, 7 voll., a cura di C. I. Gerhardt, Berlin 1875-90.

[10] Leibniz, G. W., *Monadologia e Discorso di metafisica*, Laterza, Roma-Bari 1986.

[11] Locke, J., *Saggio sull'intelletto umano*, Utet, Torino 1971.

[12] Penfield, W., *Il mistero della mente*, Vallecchi, Firenze 1991.

[13] Peruzzi, A., From Kant to Entwined Naturalism, *Annali del Dipartimento di Filosofia*, IX (1993), Università di Firenze, Olschki, Firenze 1994.

[14] Peruzzi, A., *Il significato inesistente*, F.U.P., Firenze 2004.

[15] Pescatore, G. (a cura di), *Matrix. Uno studio di caso*, Alberto Perdisa editore, Bologna 2006.

[16] Platone, *Repubblica*, Laterza, Roma-Bari 2007.

[17] Putnam, H., *Mente, linguaggio e realtà*, Mondadori, Milano 1989.

[18] Putnam, H., *Ragione, verità e storia* [1981], Mondadori, Milano 1989.

[19] Russell, B., *La filosofia di Leibniz*, Newton Compton, Roma 1972.

[20] Salucci, M., *Mente/corpo*, La Nuova Italia, Firenze 1997.

[21] Salucci, M. *La teoria dell'identità*, Le Monnier, Firenze 2005.

[22] Sani, A., *Il cinema tra storia e filosofia*, Le Lettere, Firenze 2002.

[23] Sani, A., *Il cinema pensa? Cinema, filosofia e storia*, Loescher, Torino 2008.

[24] Varela, F., Thompson, E., Rosch, E., *The Embodied Mind*, The MIT Press, Cambridge (Mass.) 1991.